

OM of: US-09-784-340-2 to: GenEmbl: * out_format : pfs
Date: Aug 27, 2001 8:11 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODE=1-frame+g2n.model -DEV=xlp
-O=/cgn2_1/USPTO.spool/US09784340/runat_27082001.123146.90/app-query.fasta.1.591
-DB=GenEmbl -QMT=fastcap -SUFFIX=olip2n.rge -GAPOP=4.500
-GAPOP=4.500 -MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000
-GAPOP=6.000 -GAPEXT=0.050 -XGAPOP=60.000 -XGAPEXT=60.000
-DELOP=6.000 -DELEXT=7.000 -XGAPOP=60.000 -XGAPEXT=60.000
-TRANS=human40.cdl -LIST=45 -DOCALLIGN=200 -THR_SCORE=quality
-THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=2000000000 -USER=US09784340 -CGEN1.1.5822 -NCPU=6
-ICPU=3 -LONGLOG -NO_XLPHY -WAIT -THREADS=1

Search information block:

Query: US-09-784-340-2

Query length: 527

Database: GenEmbl: *

Database sequences: 1344157

Database length: -856060004

Search time (sec): 1747.330000

WARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP=XGAPOP=60.000
WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000

Score list:

Sequence	Strid	Orig	ZScore	Escore	Len	Documentation
gb_p15:AK025587	+	441.00	8307.81	0.0	2786	AK025587 Homo sapiens cDNA: FI
gb_h17:AC021146	+	239.00	4457.30	4.9e-240	212904	AC021146 Homo sapiens chrom
gb_p11:AK040087	+	174.00	3259.12	2.7e-173	1650	AK040087 Sequence 10 from Patg
gb_h17:AC021146	+	75.00	1354.12	3.4e-67	212904	AC021146 Homo sapiens chrom
gb_h17:AC021146	+	42.00	731.21	1.7e-32	169246	AC021146 Homo sapiens chrom
gb_r01:AF184901	+	41.00	742.80	3.9e-33	1584	AF184901 Mus musculus olfacto
gb_p19:HSAB054	+	41.00	742.09	4.0e-33	1766	AJ006054 Homo sapiens mRNA for
gb_r02:RUGGEGNE	+	40.00	722.93	5.0e-32	1831	X57565 Rattus norvegicus mRNA
gb_p19:HSAB054	+	33.00	589.65	1.3e-24	2079	AJ006162 Homo sapiens mRNA for
gb_p14:AF081793	+	33.00	589.61	1.3e-24	2091	AF081793 Homo sapiens UDP-gluc
gb_p10:HSUPT209	+	33.00	589.60	1.3e-24	2093	Y00317 Human mRNA for liver mi
gb_p10:HSUPT209	+	33.00	589.56	1.3e-24	2107	U59209 Homo sapiens C19orf10
gb_p15:AF135416	+	33.00	572.83	1.1e-23	27310	AF135416 Homo sapiens UDP-gluc
gb_h19:AC013296	+	33.00	559.03	6.7e-23	226077	AC013296 Homo sapiens clone
gb_p15:AF112112	+	30.00	532.79	1.9e-21	2108	AF112112 Macaca fascicularis U
gb_r02:RATUDP210	+	29.00	519.46	1.1e-20	895	M35086 Rat UDP-glucuronosyltran
gb_r02:RATUDP210	+	29.00	514.34	2.1e-20	1661	M3506 Rat UDP-glucuronosyltran
gb_p15:AF294901	+	28.00	496.50	2.0e-19	1662	AF294901 Macaca mulatta UDP-gluc
gb_p14:AF016110	+	28.00	496.15	2.1e-19	1753	AF016110 Macaca fascicularis U
gb_p10:HMUDP210	+	28.00	495.78	2.2e-19	1855	U05428 Human 3,4-catechol est
gb_p10:HSUG1582	+	28.00	493.46	3.0e-19	2648	U91582 Macaca fascicularis UDP
gb_p10:HSUG1582	+	28.00	493.10	3.1e-19	2799	X63359 H. sapiens UGT2B10 mRNA
gb_h19:AC025488	+	28.00	467.59	8.3e-18	139015	AC025488 Homo sapiens chrom
gb_p15:AF294902	+	27.00	477.58	2.3e-18	1662	AF294902 Macaca mulatta UDP-gluc
gb_p15:AF112113	+	27.00	477.17	2.4e-18	1768	AF112113 Macaca fascicularis U
gb_p15:AF112113	+	27.00	453.77	4.9e-17	63705	AF112113 Homo sapiens chr13 g
gb_h19:AC025488	+	27.00	448.67	9.4e-17	139015	AC025488 Homo sapiens chrom
gb_r02:RUGGEGNE	+	24.00	420.13	3.6e-15	1844	U06273 Rattus norvegicus chro
gb_r02:RUGGEGNE	+	23.00	400.86	3.6e-15	1947	U06273 Rattus norvegicus UDP-g
gb_p14:AF072223	+	23.00	400.44	4.5e-14	2075	AF072223 Macaca fascicularis U
gb_p10:HS08854	+	23.00	400.32	4.6e-14	2090	U00854 Human UDP-glucuronosylt
gb_p15:AF180322	+	23.00	400.39	4.6e-14	2114	AF180322 Homo sapiens UDP-glu
gb_p10:HSU06641	+	23.00	400.29	4.6e-14	2123	U06641 Human UDP-glucuronosylt
gb_p13:AC019173	+	23.00	371.26	1.9e-12	181117	AC019173 Homo sapiens BAC cl
gb_r01:AF175221	+	22.00	380.18	6.1e-13	2547	AF175221 Cavia porcellus UDP-g
gb_p15:AF175221	+	21.00	363.41	5.3e-12	1832	AF175221 Oryctolagus cuniculu
gb_p14:AF016492	+	19.00	325.97	6.4e-10	1722	AF016492 Homo sapiens UDP-gluc
gb_r02:RUGGEGNE	+	19.00	325.52	6.8e-10	1846	U06274 Rattus norvegicus Sprag
gb_p14:AF064200	+	19.00	324.69	7.5e-10	2097	AF064200 Homo sapiens UDP-gluc

gb_cm:RABUGT2BC + 19.00 322.58 9.9e-10 2897 I L01083 Oryctolagus cuniculu
gb_h19:AC055794 + 19.00 297.32 2.5e-08 138665 I AC055794 Homo sapiens chr
gb_cm:RABUGT2BC + 18.00 307.25 7.1e-09 1671 I L01082 Oryctolagus cuniculu
gb_r02:RATUDP210 + 18.00 307.07 7.2e-09 1716 I Y00156 Rat mRNA for hepatic
gb_r02:RATUDP210 + 18.00 306.69 7.6e-09 1319 I M3109 Rat UDP-glucuronosyl
gb_r01:MMUDP210 + 17.00 287.63 8.7e-08 1858 I X06358 Mouse mRNA for UDP-g

seq_name: gb_p15:AK025587

seq_documentation_block:

LOCUS AK025587 2786 bp mRNA PRI 29-SEP-2000
DEFINITION Homo sapiens cDNA: FLJ21934 fls, clone HEP04364.
ACCESSION AK025587
VERSION AK025587.1 GI:10438147
KEYWORDS oligo capping, fls (full insert sequence).
SOURCE Homo sapiens hepatoma cell_line:HePg2 cDNA to mRNA, clone_1lib:HEP
clone:HEP04364.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homioidea; Homo.

REFERENCE
1 (sites)
Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project .
Unpublished (2000)
2 (bases 1 to 2786)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
Direct Submission
Submitted (29-AUG-2000) to the DDBJ/EMBL/Genbank databases. Sumio
Sugano, Institute of Medical Science, University of Tokyo.
Laboratory of Genome Structure Analysis, Human Genome Center:
Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:cdna@leims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan: cDNA full insert
sequencing. Research Association for Biotechnology: cDNA library
construction, 5'- & 3'-end one pass sequencing: Departent of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).

FEATURES

source
1..2786
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="HepG2"
/clone="HEP04364"
/note="cloning vector pME18SFL3"
71..1420
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB15179.1"
/db_xref="GI:10438148"
/translation="MPDRTENELFVDALVAVLDELSTWOSYIKNDFVIRGLTK
MNCESFYNTQTLKQLENTYDMLIDPVPICGDLMAELAVPVLTLISVGNNER
SCGLPAPLSYVPMPTGLTDMTELEKNSMLSVLFHFWIADVDHMEFYSKAL
GPTLCEVGAELIWLIRTYWDFEPQVQNFPEVGLHCKPKAKLPEKENVOS
SESDIYVSGSLFONVTEKKAIIASLADQIPKVMRGYGGKPKSTGATRLAYD
IFQNDLHPKRAPIKATVGGNACIYAIHVPYPMVQIFGQGLDINAMKKAQAE
INFKWTSDILRALRTVITDSSYENAKRSRIHDDPVKFLDRAVFTIEVMKNG
AKHLSAADLWTFQHYSDVIGFLACVAVIPLFTKCFSCCKFNKTRIERE"

COMMENT

COMMENT
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan: cDNA full insert
sequencing. Research Association for Biotechnology: cDNA library
construction, 5'- & 3'-end one pass sequencing: Departent of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).

FEATURES

source
1..2786
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="HepG2"
/clone="HEP04364"
/note="cloning vector pME18SFL3"
71..1420
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB15179.1"
/db_xref="GI:10438148"
/translation="MPDRTENELFVDALVAVLDELSTWOSYIKNDFVIRGLTK
MNCESFYNTQTLKQLENTYDMLIDPVPICGDLMAELAVPVLTLISVGNNER
SCGLPAPLSYVPMPTGLTDMTELEKNSMLSVLFHFWIADVDHMEFYSKAL
GPTLCEVGAELIWLIRTYWDFEPQVQNFPEVGLHCKPKAKLPEKENVOS
SESDIYVSGSLFONVTEKKAIIASLADQIPKVMRGYGGKPKSTGATRLAYD
IFQNDLHPKRAPIKATVGGNACIYAIHVPYPMVQIFGQGLDINAMKKAQAE
INFKWTSDILRALRTVITDSSYENAKRSRIHDDPVKFLDRAVFTIEVMKNG
AKHLSAADLWTFQHYSDVIGFLACVAVIPLFTKCFSCCKFNKTRIERE"

CDS

CDS
1..2786
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="HepG2"
/clone="HEP04364"
/note="cloning vector pME18SFL3"
71..1420
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB15179.1"
/db_xref="GI:10438148"
/translation="MPDRTENELFVDALVAVLDELSTWOSYIKNDFVIRGLTK
MNCESFYNTQTLKQLENTYDMLIDPVPICGDLMAELAVPVLTLISVGNNER
SCGLPAPLSYVPMPTGLTDMTELEKNSMLSVLFHFWIADVDHMEFYSKAL
GPTLCEVGAELIWLIRTYWDFEPQVQNFPEVGLHCKPKAKLPEKENVOS
SESDIYVSGSLFONVTEKKAIIASLADQIPKVMRGYGGKPKSTGATRLAYD
IFQNDLHPKRAPIKATVGGNACIYAIHVPYPMVQIFGQGLDINAMKKAQAE
INFKWTSDILRALRTVITDSSYENAKRSRIHDDPVKFLDRAVFTIEVMKNG
AKHLSAADLWTFQHYSDVIGFLACVAVIPLFTKCFSCCKFNKTRIERE"

BASE COUNT

BASE COUNT
919 a 463 c 496 g 908 t

alignment_scores:
Quality: 441.00 Length: 441
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

339 1LYSLYSPROSErThrLeuGlyAlaSerThrArgLeuLysAspPrIle 355
|||||
852 GAAAAAACCATCCACATTAGAGCCCACTACCTGGCTGATATTGGATA 901
|||||
356 ProGlnAsnAspLeuLeuGlyHisIspPolYsrThrLysAlaPheIleThrI 372
|||||
902 CCCGAGATGATCTTCTTGCTCATCCCAAAACCAAGCTTTATCATCTCA 951
|||||
372 sGIYGLYMetAsnGlyIleTYrGluAlaIleTYrHisGlyValrPrometY 389
|||||
952 TGGTGGAAATGATGGATGGATCTATGAAGACTATTATACATGGGGCTCAT 1001
|||||
389 aIGlyValrProIlePheGlyAspGlnLeuAspAsnIleAlaHisMetLys 405
|||||
1002 TGGGAGTTCCCATATTTGGTGTGATGAGCTTGATTAACATAGCTTACATGA 1051
|||||
406 AAlaYsGlyAlaAlaValaGluIleAsnPheLysThrMetThrSerGlnAs 422
|||||
1052 GCCAAAGAGAGAGCTGTAGAAATAACTTCAAAACCTAGACAGCGAAGA 1101
|||||
422 pLeuLeuArgAlaLeuArgThrValIleThrAspSerSerLYrGluA 439
|||||
1102 TTTTCTGAGGCTTTGAGACAGATCATTCAGGATTCCTCTATTAAGAAGA 1151
|||||
439 snAlaMetArgLeuSerArgIleHisHisAspGlnProValLysProLeu 455
|||||
1152 ATGCTATGAGATGTATCAAGAATTCACCATGATCAACCTGTAAGGCCCTA 1201
|||||
456 AspArgAlaValaPheTrpIleGluPheValaMetArgHisLYrGlyAlaLY 472
|||||
1202 GATGACACAGCTCTTCTGATCGAGTTGTCTCATGGCCCAAAAGAGCCAA 1251
|||||
472 sHisLeuArgSerAlaAlaHisAspLeuThrTriPheGlnHisTYrSerI 489
|||||
1252 GCACCTGCAGTACAGCTGCCATGCCTCACCTGGTTCAGCAGCTACTCTTA 1301
|||||
489 leAspValIleGlyPheLeuLeu 496
|||||
1302 TAGATGTGATTTGGTCTCTGCTG 1324
|||||
seq_name: gpb_hcg7:AC021146

seq_documentation_block:
LOCUS AC021146 212904 bp DNA HTG 07-JUL-2000
DEFINITION Homo sapiens chromosome 4 clone RP11-468N14, WORKING DRAFT
SEQUENCE, 24 unordered pieces.
AC021146
AC021146.4 GI:8568861
KEYWORDS
HTG, HTGS_PHASE1, HTGS_DRAFT.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
TITLE
JOURNAL
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 212904)
Waterston, R.H.
REFERENCE
AUTHORS
TITLE
JOURNAL
Direct Submission
Submitted (14-JUN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis
MO 63108, USA
COMMENT
On Jun 16, 2000 this sequence version replaced qi:7344259.

```
Sequencing vector: M13: 88%
Chemistry: Dye-terminator Big Dye; 12% of reads
Chemistry: Dye-terminator Big Dye; 12% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 199729 bases at least Q40
Consensus quality: 203731 bases at least Q30
Consensus quality: 206340 bases at least Q20
Insert size: 213000; agarose-fp
Insert size: 210604; sum-of-ctigs
Quality coverage: 3.75 in Q20 bases; agarose-fp
Quality coverage: 3.83 in Q20 bases; sum-of-ctigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1564: contig of 1564 bp in length
1565 1664: gap of unknown length
1665 3085: contig of 1421 bp in length
3086 3185: gap of unknown length
3186 5134: contig of 1949 bp in length
5135 5234: gap of unknown length
5235 8169: contig of 2935 bp in length
8170 8269: gap of unknown length
8270 10341: contig of 2072 bp in length
10342 10441: gap of unknown length
10442 13614: contig of 3173 bp in length
13615 13714: gap of unknown length
13715 17089: contig of 3375 bp in length
17090 17189: gap of unknown length
17190 20701: contig of 3512 bp in length
20702 20801: gap of unknown length
20802 25001: contig of 4200 bp in length
25002 25101: gap of unknown length
25102 29020: contig of 3919 bp in length
29021 29120: gap of unknown length
29121 33356: contig of 4236 bp in length
33357 33456: gap of unknown length
33457 36902: contig of 3446 bp in length
36903 37002: gap of unknown length
37003 42975: contig of 5973 bp in length
42976 43075: gap of unknown length
43076 48339: contig of 5264 bp in length
48340 48439: gap of unknown length
48440 56529: contig of 8090 bp in length
56530 56629: gap of unknown length
56630 65213: contig of 8584 bp in length
65214 65313: gap of unknown length
65314 74715: contig of 9402 bp in length
74716 74815: gap of unknown length
74816 88546: contig of 13731 bp in length
88547 88646: gap of unknown length
88647 10367: contig of 14721 bp in length
10368 10367: gap of unknown length
10369 118167: contig of 14700 bp in length
118168 118267: gap of unknown length
118269 132765: contig of 1498 bp in length
132766 132865: gap of unknown length
132866 154092: contig of 21227 bp in length
154093 182144: contig of 27952 bp in length
182145 182244: gap of unknown length
182245 212904: contig of 30660 bp in length

Location/Qualifiers
1..212904
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"

BASE COUNT 69110 a 36915 c 37432 g 67121 t 2326 others
ORIGIN
alignment_scores:
    Quality: 239.00      Length: 239
    Ratio: 1.000        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-784-340-2 x AC021146/rev ..

Align seg 1/1 to reverse of: AC021146 from: 1 to: 212904

1 MetArgSerSplySerAlaLeuValPheLeuLeuLeuLeuLeuPheCy 17
87413 ATGAGGCTGACAGAGTCAAGCTTGATTTCTGCTCGACGCTTCTG 87364
17 sValGlyCyGlyPheCyGlyValLeuValTTPProCyAspMetS 34
87363 TGTGGCTGTGGATTCGTGGAAAGTCGTGGCCCTGTGCACATGA 87314
34 eRIstRpleAsnValIysValIleLeuGluGluLeuIleValArgGly 50
87313 GCCATTGGCTTAATGTCAAGGTCATTTCTAGAGAGCCATAGTAGAGGC 87264
51 HisGluValThrValLeuThrHisSerIysProSerLeuIleAspTyr 67
87263 CATGAGGTAACTAGTATGACTCAAGCCCTTCGTTAATTGACTACAG 87214
67 GlyProSerAlaLeuIysPheGluValValHisMetProGluAspArg 84
87213 GAAGCTTCGTGCAATTGAATTGAGGTGCCATATGCCAGAGACAGAA 87164
84 hrcGluAsnGluIlePheValAspLeuAlaLeuAsnValLeuProGly 100
87163 CAGAAGAAATGAATATTTGTTGACCTGCTGAATGCTTGTGCCAGGC 87114
101 LeuSerThrTrpGlnSerValIleIysLeuAsnAspPheValGluI 117
87113 TTATCAACCTGGCAATCAGTTAATAAATTAATTTTGTGTGAAT 87064
117 eArgGlyThrLeuIysMetCysGluSerPheIleTyrAsnGluThrI 134
87063 AAGAGGACTTTAAAAATGATGTGTGAGACCTTATCTACAAACAGCC 87014
134 eumetIysIysLeuGluGluThrAsnTyrAspValMetLeuIleAspPro 150
87013 TTATGAAGAGCTACAGAAACCACTAGCATGTAATGCTTATAGACCT 86964
151 ValIleProCysGlyAspLeuMetAlaGluLeuLeuAlaValProPheVal 167
86963 GTGATTCCTCCGTGGAGACCTGATGGCTGAGTCTTGCAGTCCCTTTGT 86914
167 IleuThrLeuArgIleSerValIleIysAsnMetGluArgS-rcyGlyL 184
86913 GCTCACACTTAGAATTTCTGTAGAGGCAATATGAGCGCAACCTGTGGGA 86864
184 ystLeuProAlaProLeuSerTyrValProValProMetThrGlyLeuThr 200
86863 AACCTCCAGCTCCACTTCTCTAGTACCTGTGCTGATGAGAGGACATACA 86814
201 AspArgMetThrPheLeuGluValArgValIysAsnSerMetLeuSerValIle 217
86813 GACAGAAATGACCTTCTCGAAGAGTAATAAATTCATGCTTTCACAGTTT 86764
217 uphEniSphetrPIleGlnAspTyrAspTyrHisPheTrpGluIuphet 234
86763 GTTCCACTTCTGATTCAGGATTCAGCATATCTATTTTGGGAAGAGTTT 86714
234 ysrSerIysAlaLeuGly 239
|||||
```

86713 ATAGTAGCATTAAGCT 86697

seq_name: gb_pat1:AX040087

seq_documentation_block:

LOCUS AX040087 1650 bp DNA PAT 18-NOV-2000

DEFINITION Sequence 10 from Patent WO0063351.

ACCESSION AX040087

VERSION AX040087.1 GI:112330049

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1650)

AUTHORS Lal, P., Yue, H., Tang, Y. T., Hillman, J. L., Baughn, M. R. and Yang, J.

TITLE Carbohydrate-modifying enzymes

JOURNAL Patent: WO 0063351-A 10 26-OCT-2000;

Incyte Genomics, Inc. (US)

FEATURES

Source location/Qualifiers

1..1650

/organism="Homo sapiens"

/db_xref="taxon:9606"

/note="Incyte ID No: 2912330CBI"

BASE COUNT 489 a 330 c 354 g 477 t

ORIGIN

alignment_scores:

Quality: 174.00 Length: 174

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-784-340-2 x AX040087 ..

Align seg 1/1 to: AX040087 from: 1 to: 1650

354 TrrpIleProGlnAsnAspLeuGlyHisProLysThrLysAlaPheI 370

|||||

1100 TGGATACCCAGATGATCTCTTGGTCATCCCAAAACCAAGCTTTAT 1149

|||||

370 ePrrHISGLYGLYMeLAsnglyIleTyrguAlaIleTYrHISGLYValP 387

|||||

1150 CATCTATGGGGAATGATGGATCTATGAGCTATTTCAGATGGGGTCC 1199

|||||

387 rOMeValGlyValProIlePheGlyAspGlnLeuAspAsnIleAlaHis 403

|||||

1200 CTATGGTGGAGATCCCATATTGGTGCATGCTGATAACATAGCTCAC 1249

|||||

404 MeLlyAlaLysGlyAlaAlaValGluIleAsnPhelysthrMetHise 420

|||||

1250 ATGAAGGCCAAAGAGACAGCTGTAAGAAATTAACCTCAAAACTATGACAA 1299

|||||

420 rGluAspLeuAlaArgAlaLeuArgThrValIleThrAspSerSetyrl 437

|||||

1300 CGAAGATTTCAGAGGCTTGGAAACAGCATTTACCGATTCTCTTATA 1349

|||||

437 ySGLuAsnAlaMetArgLeuSerArgIleHisAspGlnProValLys 453

|||||

1350 AAGAGATGCTATGAGATTTATCAAGAAATTCACCATGATCAACTGTAAAG 1399

|||||

454 ProLeuAspArgAlaValPheTrpIleGluPheValMetArgHisLysG 470

|||||

1400 CCCCTGATGACAGCAGCTCTTGTGATGAGATTGTGTCAGCCACAAAG 1449

|||||

470 yAlaLysHisLeuArgSerAlaAlaHisAspLeuThrTrpPheGlnHis 487

|||||

1450 AGCGAAGCAGCTCGATCAGCTGCCATGACCTCAGCTGTTCCAGACT 1499

|||||

487 ySerIleAspValIleGlyPheLeuLeuThrCysValAlaThrAlaIle 503

|||||

1500 ACTCTATAGATGATGGGTTCTGCTGACCTGTGCGCAACTGCTATA 1549

504 PheLeuPheThrLysCysPheLeuPheSerCysGlnLysPheAsnLysrh 520

|||||

1550 TTCTTGTTCACAAATGTTTATTATTTCTGTCACAAATTTAATTAAC 1599

|||||

520 rArgLysIleGluLysArgGlu 527

|||||

1600 TAGAAAGATAGAAAGAGCGAA 1621

|||||

seq_name: gb_htg7:AC021146

seq_documentation_block:

LOCUS AC021146 212904 bp DNA HTG 07-JUL-2000

DEFINITION Homo sapiens chromosome 4 clone RP11-468N14, WORKING DRAFT

SEQUENCE, 24 unordered pieces.

ACCESSION AC021146

VERSION AC021146.4 GI:8568861

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 212904)

AUTHORS Waterston, R. H.

TITLE The sequence of Homo sapiens clone

JOURNAL Unpublished

2 (bases 1 to 212904)

Waterston, R. H.

Direct Submission

Submitted (14-JAN-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

On Jun 16, 2000 this sequence version replaced gi:7344259.

COMMENT

----- Genome Center -----

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc/index.shtml>

----- Project Information -----

Center project name: H_NH0468N14

----- Summary Statistics -----

Sequencing vector: M13; 88%

Sequencing vector: plasmid; 12%

Chemistry: Dye-Primer ET; 88%

Chemistry: Dye-terminator Big Dye; 12%

Assembly program: Phrap; version 0.990319

Consensus quality: 199729 bases at least Q40

Consensus quality: 203731 bases at least Q30

Consensus quality: 206340 bases at least Q20

Insert size: 213000; agarose-fp

Insert size: 210604; sum-of-contigs

Quality coverage: 3.75 in Q20 bases; pharose-fp

Quality coverage: 3.83 in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently consists of 24 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved. -----

1 1564: contig of 1564 bp in length

* 1565 1664: gap of unknown length

* 1665 3085: contig of 1421 bp in length

* 3086 3185: gap of unknown length

* 3186 5134: contig of 1949 bp in length

* 5135 5234: gap of unknown length

* 5235 8169: contig of 2935 bp in length

* 8170 8269: gap of unknown length

* 8270 10341: contig of 2072 bp in length

* 10342 10441: gap of unknown length

* 10442 13614: contig of 3173 bp in length

```

* 13615 13714: gap of unknown length
* 13715 17089: contig of 3375 bp in length
* 17090 17189: gap of unknown length
* 17190 20701: contig of 3512 bp in length
* 20702 20801: gap of unknown length
* 20802 25001: contig of 4200 bp in length
* 25002 25101: gap of unknown length
* 25102 29020: contig of 3919 bp in length
* 29021 29121: gap of unknown length
* 33356 33456: contig of 4236 bp in length
* 33457 33557: gap of unknown length
* 33558 36902: contig of 3446 bp in length
* 36903 37003: gap of unknown length
* 42975 43075: gap of 5973 bp in length
* 43076 48339: contig of 5264 bp in length
* 48340 48440: gap of unknown length
* 48441 56629: contig of 8090 bp in length
* 56630 65213: gap of unknown length
* 65214 65313: gap of 8584 bp in length
* 65314 74716: contig of 9402 bp in length
* 74717 88546: gap of unknown length
* 88547 103367: contig of 13731 bp in length
* 103368 103467: gap of unknown length
* 103468 118167: contig of 14721 bp in length
* 118168 118267: gap of unknown length
* 118268 132765: contig of 14498 bp in length
* 132766 132865: gap of unknown length
* 132866 154092: contig of 21227 bp in length
* 154093 154192: gap of unknown length
* 154193 182144: contig of 27952 bp in length
* 182145 182244: gap of unknown length
* 182245 212904: contig of 30660 bp in length
* 182245 212904: Location/Qualifiers
  source          1. 212904
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone="RP11-468N14"
                  /chromosome="4"

```

```

BASE COUNT      69110 a 36915 c 37432 g 67121 t 2326 others
ORIGIN

```

```

alignment_scores:
  quality: 75.00      Length: 75
  Ratio: 1.000        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

alignment_block:

US-09-784-340-2 x AC021146 ..

Align seg 1/1 to: AC021146 from: 1 to: 212904

```

361 leucylisprolysthrlysalaphellthrslyglymetasngl 377
|||||
33906 CTAGGTATCCCAACCAAGCTTTATCATCTCATGTGGAATGAGTGG 33955
|||||
337 YIIEYVGLAAlaIleThrsISGLyValPrometValGIyValProIleP 394
|||||
33956 GATCTATGAAGCTATTACCATGGGGTCCCTATGCTGGAGATTCCCATAT 34005
|||||
394 heglaspGINleuaspasnIlealalhismetlysalalysglYAlaAla 410
|||||
34006 TTGGTGATCAGCTTGATACATACATACATGAGGCCAAGAGCAGCT 34055
|||||
411 ValGIuIleasnPhelysthrMetThSergIuaspIleuAargAlaIle 427
|||||
34056 GTAGAAATAAATCAACTCAATATGACCAAGGAGATTTACTGAGGCTTT 34105
|||||
427 uArgThValIleThrasPserSer 435

```

```

34106 GAGAAAGTCAATACGATTCCTCG 34130
seq_name: gb_h33:AC011254

```

seq_documentation_block:

```

LOCUS      AC011254 169246 bp DNA HTG 26-MAY-2000
DEFINITION Homo sapiens chromosome 4 clone RP11-401E5 map 4, WORKING DRAFT
SEQUENCE   SEQUENCE, 12 unordered pieces.
AC011254

```

```

AC011254.3 GI:7107765
VERSION    AC011254.3
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE     human.
ORGANISM   Homo sapiens

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 169246)

Bliren,B., Linton,L., Nusbaum,C. and Lander,E.

2 (bases 1 to 169246)

Bliren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,

Baldwin,J., Barina,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B.,

Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,

Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donegan,L., Doyle,M.,

Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,

Galagan,J., Gardyna,S., Grant,G., Hagus,B., Heatford,A., Horton,L.,

Howard,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,D.,

Lehoczky,J., Lien,C., Locke,K., Macdonald,P., Margulis,N.,

McEwan,P., McGuirk,A., McKernan,K., McLaughlin,J., Meltrim,J.,

Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,

Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,

Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talmaz,J.,

Testaye,S., Tittel,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,

Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

Submitted (05-OCT-1999) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Feb 28, 2000 this sequence version replaced gi:6094610.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RN/RepeatMasker.html

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: http://www.seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L1759

Center clone name: 401.E.5

```

----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 129285 bases at least Q40
Consensus quality: 148859 bases at least Q40
Consensus quality: 160948 bases at least Q20
Insert size: 172000; agarose-fp
Quality coverage: 3.4 in Q20 bases; agarose-fp
Quality coverage: 3.5 in Q20 bases; sum-of-ctngs
-----
NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 2121: contig of 2121 bp in length
* 2122 2221: gap of 100 bp
* 2222 5198: contig of 2977 bp in length

```

```

* 5199 5298: gap of 100 bp
* 5299 10736: contig of 5438 bp in length
* 10737 10836: gap of 100 bp
* 10837 19995: contig of 9159 bp in length
* 19996 20095: gap of 100 bp
* 20096 29257: contig of 9162 bp in length
* 29258 29357: gap of 100 bp
* 29358 43327: contig of 13870 bp in length
* 43328 43327: gap of 100 bp
* 43328 53743: contig of 10416 bp in length
* 53744 53843: gap of 100 bp
* 53844 64325: contig of 10482 bp in length
* 64326 64425: gap of 100 bp
* 64426 83384: contig of 18959 bp in length
* 83385 83484: gap of 100 bp
* 83485 109710: contig of 26226 bp in length
* 109711 109810: gap of 100 bp
* 109811 134388: contig of 24578 bp in length
* 134389 134489: gap of 100 bp
* 134489 169246: contig of 34758 bp in length.
FEATURES
SOURCE
1. 169246
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-401E5"
/clone_id="RPCT-11 Human Male BAC"
1. 2121
/note="assembly-fragment"
2222. 5198
/note="assembly-fragment"
5299. 10736
/note="assembly-fragment"
10837. 19995
/note="assembly-fragment"
20096. 29257
/note="assembly-fragment"
29358. 43327
/note="assembly-fragment"
43328. 53743
/note="assembly-fragment"
vector_end:r7
vector_side:right"
misc_feature
53844. 64325
/note="assembly-fragment"
64426. 83384
/note="assembly-fragment"
83485. 109710
/note="assembly-fragment"
109811. 134388
/note="assembly-fragment"
134489. 169246
/note="assembly-fragment"
clone_end:sp6
vector_side:right"

```

BASE COUNT 52651 a 30209 c 30094 g 55185 t 1107 others
ORIGIN

alignment_scores:
Quality: 42.00 Length: 42
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-784-340-2 x AC011254 ..

Align seg 1/1 to: AC011254 from: 1 to: 169246

434 SerTyrTlysgluasnalamelargleuserarglllehlshisaspql 450
|||||TCTATTAAGAAGATGCTATGAGTTATCAAGATTACCATGATCA 16824

450 nprovallysproleuasparalavalphetpillegluphevalmet 467
|||||ACCTGTAAAGCCCTGGATGAGACGATCTCTGATGACATTCATGC 16974
467 rghstlysglyalalyshislenarg 475
|||||GCCACAAAGAGCCACACCTTCGG 16924

seq_name: gb_r01:AF184901

seq_documentation_block: 1584 bp mRNA

LOCUS AF184901 1584 bp mRNA ROD 01-OCT-2000
DEFINITION Mus musculus olfactory UDP glucuronosyltransferase mRNA, complete cds.

ACCESSION AF184901 GI:10441349
VERSION AF184901.1 GI:10441349
KEYWORDS

SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 1584)
AUTHORS Heydel,J.M.
TITLE Mus musculus mRNA for olfactory UDP glucuronosyltransferase

REFERENCE 2 (bases 1 to 1584)
AUTHORS Heydel,J.M.
TITLE Direct Submission
JOURNAL Submitted (10-SEP-1999) Biochemistry, University, Blvd Jeanne D'Arc, Dijon 21000, France

FEATURES
SOURCE location/Qualifiers
1. 1584
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
1. 1584
/codon_start=1
/product="olfactory UDP glucuronosyltransferase"
/protein_id="AA017003.1"
/db_xref="GI:10441350"

CDS

translation="MLKNILSLQSLGMSLGNVLIIMPEGSHMLNVKIIIDEL
RKHNWVIVASGALFTIPSSILFEIYVPPGKEIESVIDFVLTLENRSPS
TIMTFYEMAKVIEEPHLVSRGICDGVKINKMSLKQKEFEVLSPVPPGDIYA
LKIGIPETYSRSPASTVEKCHCKKPPRPYPAIISELTDQMSFDRVRNFTSYRA
ODYMEFELMKQMSYITRALRPITLLETMRKAEIWMKRTIWDPEPRPIYLPNEFVG
GLHCKPAKPLPKEMEERFVQISGEHGIIVFSGSVKMLTDEKAMLIASALQIPQKYL
WRVKGKIPDTLGSNTRLFDMIPQNDLGHPTKTRALITRGGTNGIYEAHGIIPMGVGP
MFADQPNIIAHMKAKGAAYEVNMTMTSSDLLNRLVYINPSYKENAMRLSRRIHDD
PVKPLDRAVEWIEEFVMRHKGAKHLRVAAHDLSPQYHSLDVTIGFLACVSAIILVAK
CCLPFKQVKCTRRKKKSD"

BASE COUNT 445 a 341 c 360 g 438 t
ORIGIN

alignment_scores:
Quality: 41.00 Length: 41
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-784-340-2 x AF184901 ..

Align seg 1/1 to: AF184901 from: 1 to: 1584

435 SerTyrTlysgluasnalamelargleuserarglllehlshisaspql 451
|||||TCTATTAAGAAGATGCTATGAGTTATCAAGATTACCATGACGAC 1303
451 ovallysproleuasparalavalphetpillegluphevalmetargh 468
|||||ACCTGTAAAGCCCTGGACGAGCTCTTGATGAGTTGTCATGCGTC 1352
1353 AGTGAAGCCCTGGACGAGCTCTTGATGAGTTGTCATGCGTC 1402

468 1slysglyalalyshtleuarg 475
|||||
1403 ACAAAGAGCCAGACCTTCGT 1425

seq_name: gb_pr9: HSA6054

seq_documentation_block:
LOCUS HSA6054 1766 bp mRNA PRI 16-JUL-1999
DEFINITION Homo sapiens mRNA for UDP glucuronosyltransferase.
ACCESSION AJ006054
VERSION AJ006054.1 GI:4753765
KEYWORDS UDP glucuronosyltransferase; ugt2A1 gene.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1766)
Jeditschky G., Cassidy A.J., Sales M., Pratt N. and Burchell B.
Cloning and characterization of a novel human olfactory
UDP-glucuronosyltransferase
Biochem. J. 340 (Pt 3), 837-843 (1999)
99289328
JOURNAL
MEDLINE
REFERENCE 2 (bases 1 to 1766)
Cassidy A.J.
Direct Submission
Submitted (07-MAY-1998) Cassidy A.J., Molecular and Cellular
Pathology, University of Dundee, University Dept. of Molecular and
Cellular pathology, Ninewells Hospital, Dundee, DD1 9SY, SCOTLAND

FEATURES
source
1. 1766
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="olfactory"
64..1647
/gene="ugt2A1"
64..1647
/gene="ugt2A1"
/EC_number="2.4.1.17"
/codon_start=1
/product="UDP glucuronosyltransferase"
/protein_id="CAB41974.1"
/db_xref="GI:4753766"
/translation="MLNNLLISLISLIGTTLGAGNYLIMPGSHMLNVKIIIDELI
KREHNTVIVASALFTTPSPSIFTEIYKVPFGKRTICVTKDPTVLENRPS
TITWYQEMAKVIEEFVLSRGICDGLKNEKMTKQREFEVLSDPVPCGDIVA
LKLGIPIYSLRSPASTVEKHCQKVPSPYPAVLSLTDMSFDRVRNISTRM
ODMFEETLMQMSYSKALGRPTLCEYMGKAEIMLRITMDPEFRPYLPNFEVG
GLCKRPKPLPKMEEFVORTSGEYGVVFLSGVMKLTTEKANLASALAOIPOKYL
WRTGKIPATLGSNTRLPDMITPONDILGHKRTAFITHGCTNGIYATIGYIPWCV
MFADQPNIAHMAKGAIVNMTMTSADLSAVRAVINPEPTKENAMLSRIHDO
PVKPLDRAVWIEFVMRKGAKHLRVANHLDFWQYHSLDVIQFLVCVTTALFLVIO
CCLEFCOKFGKIGKRRKRE"

BASE COUNT 542 a 359 c 363 g 502 t

ORIGIN

alignment_scores:
Quality: 41.00 Length: 41
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-784-340-2 x HSA6054 ..
Align seg 1/1 to: HSA6054 from: 1 to: 1766

435 SerTYLYSGluAsnAlaMetArgLeuSerArgIleHisAspGlnPr 451
|||||
1366 TCTTAAAGAGAAATGCTATGAGTTATCAAGATTCACCATGATCAAC 1415
|||||
451 oVallySProLeuAspArgAlaValPheTrpIleGluPheValMetArgH 468
|||||
1416 TGTAAAGCCCCGTGATCGAGAGCTCTTGATCGAGTTGTTCATCGGCC 1465
|||||

468 1slysglyalalyshtleuarg 475
|||||
1466 ACAAAGAGCCAGACCTTCGG 1488

seq_name: gb_r02: RRUGTGENE

seq_documentation_block:
LOCUS RRUGTGENE 1831 bp DNA ROD 14-SEP-1999
DEFINITION Rattus norvegicus mRNA UDP-glucuronosyltransferase 2A1 precursor.
ACCESSION X57565
VERSION X57565.1 GI:577762
KEYWORDS UDP-glucuronosyltransferase; ugt2A1 gene.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1831)
Lazard D., Zupko K., Porta Y., Nef P., Lazarovits J., Horn S.,
Kien M. and Lancel D.
Odorant signal termination by olfactory UDP glucuronosyl
transferase
Nature 349 (6312), 790-793 (1991)
91156050

FEATURES
source
1. 1831
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
64..1647
/gene="ugt2A1"
64..1647
/gene="ugt2A1"
/EC_number="2.4.1.17"
/codon_start=1
/product="UDP-glucuronosyltransferase 2A1 precursor"
/protein_id="CAA40797.1"
/db_xref="GI:3980217"
/translation="MLKNILLMSLISLIGMSIGAGNYLIMPGSHMLNVKIIIDELI
KREHNTVIVASALFTTPSPSIFTEIYKVPFGKRTICVTKDPTVLENRPS
TITWYQEMAKVIEEFVLSRGICDGLKNEKMTKQREFEVLSDPVPCGDIVA
LKLGIPIYSLRSPASTVEKHCQKVPSPYPAVLSLTDMSFDRVRNISTRM
ODMFEETLMQMSYSKALGRPTLCEYMGKAEIMLRITMDPEFRPYLPNFEVG
GLCKRPKPLPKMEEFVORTSGEYGVVFLSGVMKLTTEKANLASALAOIPOKYL
WRTGKIPATLGSNTRLPDMITPONDILGHKRTAFITHGCTNGIYATIGYIPWCV
MFADQPNIAHMAKGAIVNMTMTSADLSAVRAVINPEPTKENAMLSRIHDO
PVKPLDRAVWIEFVMRKGAKHLRVANHLDFWQYHSLDVIQFLVCVTTALFLVIO
CCLEFCOKFGKIGKRRKRD"

BASE COUNT 505 a 393 c 432 g 500 t 1 others

ORIGIN

alignment_scores:
Quality: 40.00 Length: 40
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-784-340-2 x RRUGTGENE ..
Align seg 1/1 to: RRUGTGENE from: 1 to: 1831

436 TYLYSGluAsnAlaMetArgLeuSerArgIleHisAspGlnProva 452
|||||
1369 TATTAAGAGAAATGCTATGAGTTATCAAGATTCACCATGATCAAC 1418
|||||
452 llySPProLeuAspArgAlaValPheTrpIleGluPheValMetArgHisL 469
|||||
1419 GAAGCCCCGTGATCGAGAGCCGTCTTGATGAGTTGTTCATGCGCACA 1468
|||||
469 ySGlyAlaLyshStleuarg 475
|||||
1469 AAGGAGCCCAAGCACCTTCGT 1488
|||||

```

seq_name: gb_pt9:HSNJ5162
seq_documentation_block: 2079 bp mRNA PRI 24-JUN-1998
LOCUS HSNJ5162 Homo sapiens mRNA for UDP-glucuronosyltransferase.
DEFINITION Homo sapiens mRNA for UDP-glucuronosyltransferase.
ACCESSION AJ005162 GI:3135024
VERSION AJ005162.1 GI:3135024
KEYWORDS UDP-glucuronosyltransferase; UGT2B4 gene.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
  1 (bases 1 to 2079)
  Ritter,J.K.
  Direct Submission
  Submitted (08-MAY-1998) Ritter J.K., Pharmacology and Toxicology,
  Virginia Commonwealth University, P.O. Box 980613, Medical College
  of Virginia, Richmond, Virginia, 23112-0613, USA
  2 (bases 1 to 2079)
  Ritter,J.K., Chen,F., Sheen,Y.Y., Lubet,R.A. and Owens,I.S.
  Two human liver cDNAs encode UDP-glucuronosyltransferases with 2
  log differences in activity toward parallel substrates including
  hydroxycholeic acid and certain estrogen derivatives
  Biochemistry 31 (13), 3409-3414 (1992)
JOURNAL
  MEDLINE
  REFERENCE
  3 (bases 1 to 2079)
  Jackson,M.R., McCarthy,L.R., Harding,D., Wilson,S., Coughtrie,M.W.
  and Burchell,B.
  Cloning of a human liver microsomal UDP-glucuronosyltransferase
  cDNA
  Biochem. J. 242 (2), 581-588 (1987)
87241362
FEATURES
  source
  1..2079
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone_lib="Lambda bacteriophage gt11"
  /tissue_lib="Liver"
  34..1620
  /gene="UGT2B4"
  34..1620
  /gene="UGT2B4"
  /EC_number="2.4.1.17"
  /function="transfers glucuronic acid to hydroxycholeic
  acid and estrogen catechols"
  /codon_start=1
  /evidence=experimental
  /product="UDP-glucuronosyltransferase"
  /protein_id="CAA06396.1"
  /db_xref="GI:3135025"
  /db_xref="SPTREMBL:O60867"
  /translation="MSKMTSALLLIQSCYFSSGCGKVLWPIESHMNIKITLD
  ELVORGEVTVLASASTISDPNSPSITKEFYVYVYVSELSQWTFLEKVMELPK
  DTFMSYFQVOEIMATFNDILRKFCIDIVSKMKLKQESRFDFVLDAVFPFELP
  AELKIPFVYSLRSPGYAIEKSGGLFPPSYVYVSELSQWTFLEKVMELPK
  YEEFMEOLFDMKKMDQFSEVIGRPTTISLMAKADIVLINYMDQFPHLLPVEE
  VGGIHCRAKRLPKMEEFNOSGNGVYVSLGMSVNTSEERANVLSAKATPQK
  VLMRDKKRPDTGLNTRLYKWIYFNDLLGPRKTAFTTHGGANGIYEAIGHLPWG
  VPLRADPDNTAHMKAGAAVSLDFHMSSTDNLNALTIVNDIPYKKNAMKLSIHH
  DQPKPLDRAVFIETFVIRHKAGNHLRVAADHLTWFOYHSLDVIGFLACVATVFI
  TKLFCVWKFEVRTGKKGRD"
  546
  /gene="UGT2B4"
  /citation=[3]
  /replace="g"
  548..549
  /gene="UGT2B4"
  /citation=[3]
  /replace="ct"
  1187..1188
  /gene="UGT2B4"
  conflict
  conflict
  conflict
  conflict

```

```

/citation=[3]
/replace="ct"
1190
/gene="UGT2B4"
/citation=[3]
/replace="c"
1192
/gene="UGT2B4"
/citation=[3]
/replace="a"
1977..1978
/citation=[3]
/replace="cga"
2068
/citation=[3]
/replace="g"
BASE COUNT 634 a 396 c 436 g 613 t
ORIGIN
alignment_scores:
  Quality: 33.00 Length: 33
  Ratio: 1.000 Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
  US-09-784-340-2 x HSNJ5162
Align seq 1/1 to: HSNJ5162 from: 1 to: 2079
443 leuSerArgIleHisHisAspGlnProValIysProLeuAspArgAlaVala 459
|||||
1366 TTTATCAGAAATTCATCATGATCAACGATGAGCCCTTATCATGACAGT 1415
459 lphETPRLcGlupheValMetargHisLysGlyAlaLysHisLeuArg 475
|||||
1416 CTTCGTGATTGAATTGTCTCATGCGCCATTAAGGAGCCACACCTTCGG 1464
seq_name: gb_pt4:AF081793
seq_documentation_block:
LOCUS AF081793 2091 bp mRNA PRI 18-AUG-1998
DEFINITION Homo sapiens UDP glucuronosyltransferase 2B4 precursor (UGT2B4)
ACCESSION AF081793
VERSION AF081793.1 GI:3426331
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
  1 (bases 1 to 2091)
  Jin,C.J., Miners,J.O., Lillywhite,K.J. and Mackenzie,P.I.
  cDNA cloning and expression of two new members of the human liver
  UDP-glucuronosyltransferase 2B subfamily
  Biochem. Biophys. Res. Commun. 194 (1), 496-503 (1993)
JOURNAL
  MEDLINE
  REFERENCE
  2 (bases 1 to 2091)
  Mackenzie,P.I.
  Direct Submission
  Submitted (30-JUL-1998) Clinical Pharmacology, Flinders University
  of South Australia, Bedford Park, SA 5042, Australia
  1..2091
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /chromosome="4"
  1..2091
  /gene="UGT2B4"
  /note="Formerly UGT2B11"
  /allele="UGT2B4*L109, L396 variant"
  36..1622
  /gene="UGT2B4"
  conflict
  conflict
  conflict
  conflict

```



```
/codon_start=1
/product="UDP-glucuronosyltransferase 2B4 precursor"
/protein_id="AAC32272.1"
/db_xref="GI:3426332"
/translation="MSKMTSALLLIQLSCYSSGCGKVLVPTPESHMMNITLID
ELVORGEHTVLAASSASISFDPNPSSTLKEVYPSLITKEFEDIIKOLVRAAEILPK
DTPWYSQVOEIMTFENDILRKFCDIVSNKMLKLOESRDVVLADAVPEFGELL
AELKIPDYVRRPFSFGYALIEKHSGLLEPPSYVPMVSELSQDMFIERVKMIVL
YEFEMFOIPMKKMOQYSEVIGRPTLTSTMAKADIMLIRNYWDOFPHLLPNVEF
VGLHCKPKPLPKKKEFVQSSGNGVYVFSIGSVNSNTSEERANVIALAKIPQK
VLMRFGNKPDTLGLNTRLYKWLIPQNDLGHPTKRAFTTHGANGIYEAHYH1PMVG
VPLADQPNIAHMAKGAAVSLDFHTMSSTDLLNLKTVINDPLVKENMKSRIHH
DQVPLDRAVVEIEFVMRHKGAHLRVAHADLTPWQYHSLDVTGFLACVATVIFI
TKCLFCVMKFPVRFGKKGRD"
BASE COUNT      638 a      401 c      437 g      615 t
ORIGIN

alignment_scores:
  Quality:      33.00      Length:      33
  Ratio:        1.000      Gaps:      0
  Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
  US-09-784-340-2 x AF081793 ..
  Align seg 1/1 to: AF081793 from: 1 to: 2091

443 LeuSerArgIleHisAspGlnProValLysProLeuAspArgAlaVala 459
|||||
1368 TTATCAGAAATTCATCATGATCAACAGCAGCCCTTGATCGAGCAGT 1417
459 lPheTrpIleGluPheValMetArgHisLysGlyAlaLysHisLeuArg 475
|||||
1418 CTCTGGATTGAATTTGTCATGCCCATTAAGACCAAGCACCTTCGG 1466
seq_name: gb_pr10:HSUDPCT

seq_documentation_block:
LOCUS      HSUDPCT      2093 bp      mRNA      PRI      23-MAR-1995
DEFINITION Human mRNA for liver microsomal UDP-glucuronosyltransferase
              (UDPCT).
ACCESSION   Y00317
VERSION     Y00317.1 GI:37588
KEYWORDS    transferase; UDP-glucuronosyltransferase.
SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 2093)
AUTHORS     Burchell, B.
TITLE       Direct Submission
JOURNAL     Submitted (26-JAN-1987) Burchell B., Dept. Biochemistry, Medical
2 (bases 1 to 2093)
AUTHORS     Jackson, M.R., McCarthy, L.R., Harding, D., Wilson, S., Coughtrie, M.W.
and Burchell, B.
TITLE       Cloning of a human liver microsomal UDP-glucuronosyltransferase
CDNA
JOURNAL     Biochem. J. 242 (2), 581-588 (1987)
MEDLINE     87241362
FEATURES
  source
    1..2093      Location/Qualifiers
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /tissue_type="liver"
    /clone_lib="lambda gt11 liver cDNA"
    38..106
    /product="put. signal peptide"
    38..1624
    /note="precursor"
    /codon_start=1
```

```
/protein_id="CAA68415.1"
/db_xref="GI:37589"
/db_xref="SWISS-PROT:P06133"
/translation="MSKMTSALLLIQLSCYSSGCGKVLVPTPESHMMNITLID
ELVORGEHTVLAASSASISFDPNPSSTLKEVYPSLITKEFEDIIKOLVRAAEILPK
DTPWYSQVOEIMTFENDILRKFCDIVSNKMLKLOESRDVVLADAVPEFGELL
AELKIPDYVRRPFSFGYALIEKHSGLLEPPSYVPMVSELSQDMFIERVKMIVL
YEFEMFOIPMKKMOQYSEVIGRPTLTSTMAKADIMLIRNYWDOFPHLLPNVEF
VGLHCKPKPLPKKKEFVQSSGNGVYVFSIGSVNSNTSEERANVIALAKIPQK
VLMRFGNKPDTLGLNTRLYKWLIPQNDLGHPTKRAFTTHGANGIYEAHYH1PMVG
VPLADQPNIAHMAKGAAVSLDFHTMSSTDLLNLKTVINDPLVKENMKSRIHH
DQVPLDRAVVEIEFVMRHKGAHLRVAHADLTPWQYHSLDVTGFLACVATVIFI
TKCLFCVMKFPVRFGKKGRD"
BASE COUNT      638 a      397 c      439 g      619 t
ORIGIN

mat_peptide
misc_feature
misc_feature
misc_feature
BASE COUNT      638 a      397 c      439 g      619 t
ORIGIN

alignment_scores:
  Quality:      33.00      Length:      33
  Ratio:        1.000      Gaps:      0
  Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
  US-09-784-340-2 x HSUDPCT ..
  Align seg 1/1 to: HSUDPCT from: 1 to: 2093

443 LeuSerArgIleHisAspGlnProValLysProLeuAspArgAlaVala 459
|||||
1370 TTATCAGAAATTCATCATGATCAACAGCAGCCCTTGATCGAGCAGT 1419
459 lPheTrpIleGluPheValMetArgHisLysGlyAlaLysHisLeuArg 475
|||||
1420 CTCTGGATTGAATTTGTCATGCCCATTAAGACCAAGCACCTTCGG 1468
seq_name: gb_pr10:HSU59209

seq_documentation_block:
LOCUS      HSU59209      2107 bp      mRNA      PRI      02-JUL-1998
DEFINITION Homo sapiens C19steroid specific UDP-glucuronosyltransferase mRNA,
              complete cds.
ACCESSION   U59209
VERSION     U59209.1 GI:3287472
KEYWORDS    complete cds.
SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 2107)
AUTHORS     Beaulieu, M., Levesque, E., Hum, D.W. and Belanger, A.
TITLE       Isolation and characterization of a novel cDNA encoding a human
UDP-glucuronosyltransferase active on C19 steroids
JOURNAL     J. Biol. Chem. 271 (37), 22855-22862 (1996)
MEDLINE     96394358
REFERENCE   2 (bases 1 to 2107)
AUTHORS     Hum, D.W., Belanger, A., Beaulieu, M. and Levesque, E.
TITLE       Direct Submission
JOURNAL     Submitted (23-MAY-1996) Laboratory of Molecular Endocrinology,
Centre Hospitalier de l'universite Laval, 2705 Boul. Laurier,
Ste-Foy, Quebec G1V 4G2, Canada
CDNA
JOURNAL
MEDLINE
FEATURES
  source
    1..2107      Location/Qualifiers
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /tissue_type="prostate"
    /cell_type="LNCaP"
    52..1644
```

```

/note="UGT2B17"
/codon_start=1
/product="C19steroid specific UDP-glucuronosyltransferase"
/protein_id="AAC2549.1"
/db_xref="GI:3287473"
/translation="MSLKMWSVFLIMOLSCYFSSGCGKYLWPTFESHMINKTILE
ELVORGEHVIVLSSASILVNASKSAIKLEVPTSLTKNDLEDFMKMDPMYTSIS
KNFWSYFQLOELCMEYSDYNIKCEDAVLNMKMKLOESKEDVLADAVNCGEL
LAELINIPFLYSLEFSGYTYEKNGGFLPPSPYVPMSELSDOMTPEMRKMTYM
LYEDFMFOADYDKMDQFSEVIGRPTLLETMGKAMMTLITRYMDEFPRLPND
FVGLHCKPAKPLPKEMEEFVQSSGNGIYVFSIGSISNMSESNAMTISALQITPO
KULMRGCKRPNTLGSNTRIKMLPQNDLIGHPTKAFITHGNGIYEAIGHIPIYV
GILFADQHDNIAMKAKGALSVDLITMSSRDILNALKSVINDPIYKENIMKLSRTH
HDQVPLDRAVWIEFVMRHKGAKHLRVAHNLWLTQYHSLDVIATLACVATMIRM
ITKCLFCFERRKAKTKGKKRD"
BASE COUNT      657 a      382 c      424 g      644 t
ORIGIN
alignment_scores:
  Quality:      33.00      Length:      33
  Ratio:        1.000      Gaps:      0
Percent Similarity: 100.000      Percent Identity: 100.000
alignment_block:
US-09-784-340-2 x HSUS9209
Align seg 1/1 to: HSUS9209 from: 1 to: 2107
443 LeuserArgIIeHIShISpGlnProValIysProLeuAspArgAlaVala 459
|||||
1367 TTATCAAGAAATTCATCATGATCAACCGGTGAAGCCCTGGATCGACGAGT 1436
459 IPhetRPIlegIupheValMeIarGHIslYsgIYAlaIYslHleuArg 475
|||||
1437 CTCTCGATGTAGTTGTCTCATGCCCATTAAGACCAAGCAACCTTCGG 1485
seq_name: gb-pr5:AF135416
seq_documentation_block:
LOCUS      AF135416      27310 bp      DNA
DEFINITION Homo sapiens UDP-glucuronosyltransferase gene, complete cds.
ACCESSION  AF135416
VERSION    AF135416.1 GI:8650277
KEYWORDS
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE  1 (bases 1 to 27310)
AUTHORS   Riedy,M. and Miller,A.
TITLE     Genomic organization and structure of the UGT2B gene complex at
JOURNAL   unpublished
REFERENCE  2 (bases 1 to 27310)
AUTHORS   Riedy,M. and Miller,A.
TITLE     Direct Submission
JOURNAL   Submitted (18-MAR-1999) Axyx Pharmaceuticals, 11099 N Torrey Pines
          Rd., La Jolla, CA 92039, USA
FEATURES
  source
  location/Qualifiers
    1..27310
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /chromosome="4"
    /map="4q13"
    /clone="Research Genetics BAC 321d14"
    complement(join(<1169..1445,5738..5957,7139..7226,
9969..10100,14223..14371,15671..>16391))
/product="UDP-glucuronosyltransferase"
complement(join(1169..1445,5738..5957,7139..7226,
9969..10100,14223..14371,15671..16391))
/codon_start=1
/product="UDP-glucuronosyltransferase"

```

```

/protein_id="AAF78145.1"
/db_xref="GI:8650278"
/translation="MSMKWTSAALLIOLSCYFSSGCGKYLWPTFESHMINKTILD
ELVORGEHVIVLSSASISFPDNPSPSTLKEEYVSLTKNEEDIIKOLVRMAELK
DTEWSYFQLOELCMEYSDYNIKCEDAVLNMKMKLOESKEDVLADAVNCGEL
LAELINIPFLYSLEFSGYTYEKNGGFLPPSPYVPMSELSDOMTPEMRKMTYM
LYEDFMFOADYDKMDQFSEVIGRPTLLETMGKAMMTLITRYMDEFPRLPND
FVGLHCKPAKPLPKEMEEFVQSSGNGIYVFSIGSISNMSESNAMTISALQITPO
KULMRGCKRPNTLGSNTRIKMLPQNDLIGHPTKAFITHGNGIYEAIGHIPIYV
GILFADQHDNIAMKAKGALSVDLITMSSRDILNALKSVINDPIYKENIMKLSRTH
HDQVPLDRAVWIEFVMRHKGAKHLRVAHNLWLTQYHSLDVIATLACVATMIRM
ITKCLFCFERRKAKTKGKKRD"
BASE COUNT      8660 a      5083 c      4901 g      8466 t
ORIGIN
alignment_scores:
  Quality:      33.00      Length:      33
  Ratio:        1.000      Gaps:      0
Percent Similarity: 100.000      Percent Identity: 100.000
alignment_block:
US-09-784-340-2 x AF135416/rev
Align seg 1/1 to reverse of: AF135416 from: 1 to: 27310
443 LeuserArgIIeHIShISpGlnProValIysProLeuAspArgAlaVala 459
|||||
1423 TTATCAAGAAATTCATCATGATCAACCGGTGAAGCCCTGGATCGACGAGT 1374
459 IPhetRPIlegIupheValMeIarGHIslYsgIYAlaIYslHleuArg 475
|||||
1373 CTCTCGATGTAGTTGTCTCATGCCCATTAAGACCAAGCAACCTTCGG 1325
seq_name: gb.htg4:AC013296
seq_documentation_block:
LOCUS      AC013296      226077 bp      DNA      HTG      13-JUL-2000
DEFINITION Homo sapiens clone RP11-3M18, LOW-PASS SEQUENCE SAMPLING.
ACCESSION  AC013296
VERSION    AC013296.5 GI:9121251
KEYWORDS   HTG; HTGS_PHASE0.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE  1 (bases 1 to 226077)
AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE     Homo sapiens, clone RP11-3M18
JOURNAL   unpublished
REFERENCE  2 (bases 1 to 226077)
AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
          Batwin,J., Barna,N., Beckerly,R., Boguslavskiy,L., Boukhallier,B.,
          Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
          Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donnell,L., Doyle,M.,
          Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,
          Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
          Howland,J.C., Johnson,R., Jones,C., Kann,L., Karates,A., Klein,J.,
          Lehoczy,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
          McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
          Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
          Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
          Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
          Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
          Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
JOURNAL   Direct Submission
TITLE     Submitted (06-NOV-1999) Whitehead Institute/MIT Center for Genome
          Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT   On Jul 13, 2000 this sequence version replaced gi:6514003.
          All repeats were identified using RepeatMasker:
          Smit, A.F.A. & Green, P. (1996-1997)
          http://ftp.genome.washington.edu/RM/RepeatMasker.html
          ----- Genome Center

```

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information -----
Center project name: L2775
Center clone name: 3_M18

* NOTE: This record contains 246 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overall relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 814: contig of 814 bp in length
* 815 914: gap of 100 bp
* 915 1704: contig of 790 bp in length
* 1705 1804: gap of 100 bp
* 1805 2602: contig of 798 bp in length
* 2603 2702: gap of 100 bp
* 2703 3484: contig of 782 bp in length
* 3485 3584: gap of 100 bp
* 3585 4378: contig of 794 bp in length
* 4379 4478: gap of 100 bp
* 4479 5257: contig of 779 bp in length
* 5258 5357: gap of 100 bp
* 5358 6150: contig of 793 bp in length
* 6151 6250: gap of 100 bp
* 6251 7038: contig of 788 bp in length
* 7039 7138: gap of 100 bp
* 7139 7940: contig of 802 bp in length
* 7941 8040: gap of 100 bp
* 8041 8869: contig of 829 bp in length
* 8870 8969: gap of 100 bp
* 8970 9790: contig of 821 bp in length
* 9791 9890: gap of 100 bp
* 9891 10695: contig of 805 bp in length
* 10696 10795: gap of 100 bp
* 10796 11592: contig of 797 bp in length
* 11593 11692: gap of 100 bp
* 11693 12475: contig of 783 bp in length
* 12476 12575: gap of 100 bp
* 12576 13364: contig of 789 bp in length
* 13365 13464: gap of 100 bp
* 13465 14255: contig of 791 bp in length
* 14256 14355: gap of 100 bp
* 14356 15142: contig of 787 bp in length
* 15143 15242: gap of 100 bp
* 15243 16023: contig of 781 bp in length
* 16024 16123: gap of 100 bp
* 16124 16932: contig of 809 bp in length
* 16933 17032: gap of 100 bp
* 17033 17878: contig of 846 bp in length
* 17879 17978: gap of 100 bp
* 17979 18786: contig of 808 bp in length
* 18787 18886: gap of 100 bp
* 18887 19677: contig of 791 bp in length
* 19678 19777: gap of 100 bp
* 19778 20575: contig of 798 bp in length
* 20576 20675: gap of 100 bp
* 20676 21463: contig of 788 bp in length
* 21464 21563: gap of 100 bp
* 21564 22350: contig of 787 bp in length
* 22351 22450: gap of 100 bp
* 22451 23220: contig of 770 bp in length
* 23221 23320: gap of 100 bp
* 23321 24108: contig of 788 bp in length
* 24109 24208: gap of 100 bp

24209 24969: contig of 761 bp in length
* 24970 25069: gap of 100 bp
* 25070 25852: contig of 783 bp in length
* 25853 25952: gap of 100 bp
* 25953 26726: contig of 774 bp in length
* 26727 26826: gap of 100 bp
* 26827 27625: contig of 799 bp in length
* 27626 27725: gap of 100 bp
* 27726 28540: contig of 815 bp in length
* 28541 28640: gap of 100 bp
* 28641 29413: contig of 773 bp in length
* 29414 29513: gap of 100 bp
* 29514 30314: contig of 801 bp in length
* 30315 30414: gap of 100 bp
* 30415 31215: contig of 801 bp in length
* 31216 31315: gap of 100 bp
* 31316 32099: contig of 784 bp in length
* 32100 32199: gap of 100 bp
* 32200 32995: contig of 796 bp in length
* 32996 33095: gap of 100 bp
* 33096 33892: contig of 797 bp in length
* 33893 33992: gap of 100 bp
* 33993 34782: contig of 790 bp in length
* 34783 34882: gap of 100 bp
* 34883 35658: contig of 776 bp in length
* 35659 35758: gap of 100 bp
* 35759 36554: contig of 796 bp in length
* 36555 36654: gap of 100 bp
* 36655 37443: contig of 789 bp in length
* 37444 37543: gap of 100 bp
* 37544 38361: contig of 818 bp in length
* 38362 38461: gap of 100 bp
* 38462 39228: contig of 767 bp in length
* 39229 39328: gap of 100 bp
* 39329 40157: contig of 829 bp in length
* 40158 40257: gap of 100 bp
* 40258 41059: contig of 802 bp in length
* 41060 41159: gap of 100 bp
* 41160 41959: contig of 800 bp in length
* 41960 42059: gap of 100 bp
* 42060 42830: contig of 771 bp in length
* 42831 42930: gap of 100 bp
* 42931 43733: contig of 803 bp in length
* 43734 43833: gap of 100 bp
* 43834 44657: contig of 824 bp in length
* 44658 44757: gap of 100 bp
* 44758 44554: contig of 797 bp in length
* 44555 45654: gap of 100 bp
* 45655 46440: contig of 786 bp in length
* 46441 46540: gap of 100 bp
* 46541 47356: contig of 816 bp in length
* 47357 47456: gap of 100 bp
* 47457 48260: contig of 804 bp in length
* 48261 48360: gap of 100 bp
* 48361 49165: contig of 805 bp in length
* 49166 49265: gap of 100 bp
* 49266 50058: contig of 793 bp in length
* 50059 50158: gap of 100 bp
* 50159 50968: contig of 810 bp in length
* 50969 51068: gap of 100 bp
* 51069 51841: contig of 773 bp in length
* 51842 51941: gap of 100 bp
* 51942 52748: contig of 807 bp in length
* 52749 52848: gap of 100 bp
* 52849 53670: contig of 822 bp in length
* 53671 53770: gap of 100 bp
* 53771 54554: contig of 784 bp in length
* 54555 54654: gap of 100 bp
* 54655 55442: contig of 788 bp in length
* 55443 55542: gap of 100 bp
* 55543 56341: contig of 799 bp in length
* 56342 56441: gap of 100 bp
* 56442 57261: contig of 820 bp in length

```

* 57262 57361: gap of 100 bp
* 57362 58170: contig of 809 bp in length
* 58171 58270: gap of 100 bp
* 58271 59048: contig of 778 bp in length
* 59049 59148: gap of 100 bp
* 59149 59954: contig of 806 bp in length
* 59955 60054: gap of 100 bp
* 60055 60858: contig of 804 bp in length
* 60859 60958: gap of 100 bp
* 60959 61761: contig of 803 bp in length
* 61762 61861: gap of 100 bp
* 61862 62631: contig of 770 bp in length
* 62632 62731: gap of 100 bp
* 62732 63524: contig of 793 bp in length
* 63525 63624: gap of 100 bp
* 63625 64411: contig of 787 bp in length
* 64412 64511: gap of 100 bp
* 64512 65269: contig of 758 bp in length

```

alignment_scores:

```

Quality: 33.00 Length: 33
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

alignment_block:

US-09-784-340-2 x AC013296 ..

Align seg 1/1 to: AC013296 from: 1 to: 226077

443 Leuserarlg1ehshsaspGlnProvalysProleuaspargalaya 459

178261 TTATCAAGAAATTCATCATGATCAACCAAGCCCTTGATCGACAGC 178310

459 lPheTrp1leGluPheValMetArgHisLysGlyAlaLysHisLeuArg 475

178311 CTCTGATTCATATTGTGATGCGCCATTAAGGACCAAGCACCCTTCGG 178359

seq_name: gb_pfs:AF112112

seq_documentation_block:

```

LOCUS AF112112 2108 bp mRNA PRI 11-APR-1999
DEFINITION Macaca fascicularis UDP-glucuronosyltransferase 2B19 precursor,
mRNA, complete cds.
ACCESSION AF112112
VERSION AF112112.1 GI:4580601
KEYWORDS
SOURCE
ORGANISM
. crab-eating macaque.
Macaca fascicularis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
REFERENCE
AUTHORS Belanger,G., Barbier,O., Hum,D.W. and Belanger,A.
TITLE Molecular cloning, expression and characterization of a monkey
steroid UDP-glucuronosyltransferase, UGT2B19, that conjugates
testosterone
JOURNAL Eur. J. Biochem. 260 (3), 701-708 (1999)
MEDLINE 99203465
REFERENCE
AUTHORS Belanger,G., Barbier,O., Hum,D.W. and Belanger,A.
TITLE Direct Submission
JOURNAL Submitted (07-DEC-1998) Molecular Endocrinology, CHUL Research
Center, 2705 Laurier Blvd, Sainte Foy, Quebec G1V 4G2, Canada
FEATURES
SOURCE
location/Qualifiers
1..2108
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
26..1612
/EC_number="2.4.1.17"
/note="UGT2B19"
/codon_start=1
/product="UDP-glucuronosyltransferase 2B19 precursor"
/protein_id="AAD24435.1"

```

```

/db_xref="GI:4580602"
/translation="MSMKWTSAALLLQLSCYLSFGSGKVLVMPTEFSHMNIKITLID
ELVORHEVTVLAVSTSLPEPNNSPLKFEICPTSLTEFQDSVTOLVSRSDIRK
DTFWPHLEHVOEMMTYIGDMIRKFCQDVSNKKLMKLOESRQDVLAQASPCGELL
AELKLPFYVSLRSPGYALEKHGGELFPPSYVPVTMSLRDQMTFMRNOMITMV
YEPFMFOVMDVKRMDOEYSKVCGRPTTLFEINAKATINLIRNMDQFFHPLPNEF
VGLSHCKPAKPLPEKMEEFYVQSSGNGVVFSLGSVSNSEERANVIALAKITPK
VLMRFQGNRPDTIGLNTQYKWLPPQNDLGHPRTRAFITHGANGIYEA1YHGIPMG
VPLEADQPNIAIMKAKAAGVRLDFDMSSTDLNALKTVINDPIYKENAMKLSYIH
DQVPEKLDRAVEMIEFVNRHKGAHLRVAADLTWQYHSLDIVIGFLACVATVFII
TKCLFCVMKFEVTRKKGRD"
BASE COUNT 649 a 393 c 438 g 628 t
ORIGIN

```

alignment_scores:

```

Quality: 30.00 Length: 30
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

alignment_block:

US-09-784-340-2 x AF112112 ..

Align seg 1/1 to: AF112112 from: 1 to: 2108

446 lLeHSHsAspGlnProvalysProleuaspargalaya1PheTrp1l 462

1367 ATTGATCATGATGATCAACCAAGCCCTTGATCGACAGCTTCTGGAT 1416

462 eGluPheValMetArgHisLysGlyAlaLysHisLeuArg 475

1417 TGAATTGTGATGCGCCATTAAGGACCAACACCTTCGG 1456

THIS PAGE BLANK (USPTO)